

Colab Guide for Beginners

Week 2 Additional · Worked on the reading's data

What This Document Covers

You do not need to know Python. Google Colab has an AI assistant, Gemini, built in. You give it data and ask questions in plain English, and it writes and runs the code for you. Your job is to ask good questions and check the answers. This guide walks through one full example, building the four scatter plots in the reading and fitting the line, the log-log curve, and the U-shape. When you are done, open `2_Bonus_Question.pdf` in this package and try it on new data.

Talk to Gemini, Step by Step

File you need. `data/example_four_relationships.csv` in this package. It holds the four pairs from the reading, one row per period. Each row carries the pair, the name and unit of X , the name and unit of demand, and the source.

Set up, about two minutes

- (1) Go to colab.research.google.com, sign in with any Google account, and choose *File, New notebook*.
- (2) Click the **Gemini** button, a small spark icon. A chat panel opens on the right. On a blank notebook you may also see a link such as *Analyze files with Gemini*, which opens the same panel.
- (3) Upload the CSV in the chat panel. Look for an *add files* or **+** button. If you cannot find it, drag the file into the **folder icon** panel on the left, and name the file in your first message.

Colab changes its buttons from time to time, so the labels you see may differ slightly. The steps stay the same.

Have the conversation

Send the turns below one at a time. Read each answer before you send the next. When Gemini offers code, click the button that adds and runs it, which may say *Accept and run* or *Execute plan*. You do not need to read the code. You do need to read the results.

Turn 1. Understand the file before any analysis

I uploaded `example_four_relationships.csv`. Before you analyze anything, tell me what is in the file: the four pairs, what X and demand are in each pair with their units, the first and last period, the number of rows per pair, and any gaps. Do not compute anything else yet.

Check. Gemini should find four pairs: 92 months of mortgage rates and home sales, 129 months of Wikipedia views, 169 weeks of avocado prices and volumes, and 731 days of temperature and bike rides. The avocado weeks end on Sundays. If it reports a gap that is not there, ask it to show the rows.

Turn 2. Draw the dots and measure how tightly they hug a line

Draw a 2 by 2 figure with one scatter plot per pair, one dot per period, X on the horizontal axis and demand on the vertical axis. Put the correlation r in each panel title, and label both axes with units. For the avocado pair, use logarithmic axes. Then tell me, for each pair, whether the dots look like a straight line, a curve that keeps bending one way, or a curve that rises and then falls.

Check. Compare with the reading. Mortgage rates and home sales, r about -0.57 . Air conditioning and barbecue, about 0.82 . Avocado price and volume, about -0.5 . Temperature and rides, about 0.63 . The bike panel should bend over at the warm end.

Turn 3. Fit the shape

For the first two pairs, fit a straight line and tell me the slope in its units: homes per point of mortgage rate, and barbecue views per air-conditioning view. For the avocado pair, fit a straight line to the logarithms of price and volume and tell me the slope; that slope is the elasticity. For the bike pair, fit a curve with temperature and temperature squared, and tell me the temperature at which the fitted curve peaks. Show the arithmetic for one example of each.

Check. About -34 thousand homes a year per point of rate. About 0.30 barbecue views per air-conditioning view. An avocado elasticity near -0.6 , so a 10% higher price goes with about 6% less volume. A bike peak near 29 degrees Celsius, with the slope positive below it and negative above it.

Turn 4. Ask the three questions, and separate fact from interpretation

For each pair, answer three questions in one sentence each. Could something else move both X and demand at the same time? Could the arrow run from demand to X instead? Are there cases or parts of the range we cannot see in these data? Label each sentence as Data if it comes from the file, or Interpretation if it is your reading.

Check. Every sentence labeled Data should match a number from Turn 2 or Turn 3. The summer, the Super Bowl, a short crop, and a mortgage rate buydown are interpretations. They are good ones, but the file does not contain them.

Turn 5. Push back

Which of the four lines would you pull as a lever, and which would you only forecast with? For the bike curve, what happens to the peak temperature if you drop all days above 30 degrees? For the avocado elasticity, what happens if you fit each year separately?

Check. Disagree with Gemini if you think it is wrong. That disagreement is good material for your share.

Using your own data

Download your pair as a CSV with one row per period and two columns of numbers, upload it, and open with this turn. It makes Gemini ask before it computes.

Opening turn for your own pair

I uploaded [file name]. It has [what X measures] and [what demand measures] from [source], one row per [period]. Before you compute anything, ask me the questions you need to understand what could move these two series. Then draw them as a scatter plot, fit a line, and tell me whether a log-log curve or a U-shape fits better than the line.

Good free sources with a download button include [FRED](#) for U.S. economic data, [Our World in](#)

Data, the [World Bank](#), the [EIA](#) for energy, and the [Wikipedia pageviews tool](#) for attention to a topic.

Three rules for talking to any AI about data

- (1) Give it the data. Never ask it to recall numbers from memory.
- (2) Ask it to show one worked example, so you can check the arithmetic.
- (3) Read the results, not the code. If a number looks wrong, say so in plain words and ask it to recheck.

If Something Goes Wrong

- **Red error text.** Tell Gemini in plain words that the code gave an error and ask it to fix and rerun it.
- **An error that mentions a website or a code such as 503.** A data source did not answer. Wait a minute and try again.
- **Nothing runs.** If the top right corner says *Connect*, click it and wait, then run again.
- **A number does not match the reading.** Ask Gemini to show the rows it used and the arithmetic. Sources are revised over time, so a small difference can be a revision. A large one means something went wrong.
- **The log-log fit fails.** A zero or a negative value has no logarithm. Ask Gemini to show any rows with zero volume and leave them out.